nature research

Corresponding author(s): Kenneth Zaret, Makiko Iwafuchi

Last updated by author(s): 1/18/2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Nikon NIS Elements was used to collect imaging data
Data analysis	RNA-seq and ATAC-seq alignment: STAR-2.5.2a RNA-seq read count: HTSeq-0.6.1 RNA-seq normalization: DESeq2 Gene ontology enrichment analyses: DAVID Bioinformatics Resources 6.8 ATAC-seq peak calling: MACS2 novo motif analysis: HOMER-v4.9 ATAC-seq peak intersection: BEDtools v2.15.0 CLUSTAL OMEGA (EBI online version: https://www.ebi.ac.uk/Tools/msa/clustalo/)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The ATAC-seq and RNA-seq data have been deposited in the Gene Expression Omnibus database under accession number GSE134465.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For high-throughput sequencing experiments, two (ATAC-seq) or three (RNA-seq) independent biological replicates were used. Sample sizes for other assays were not predetermined and were chosen based on our prior experience and common standards in the field for detecting statistically significant differences between conditions.
Data exclusions	No data were excluded from the analysis
Replication	Attempts at replication for the sequencing study were successful, verified by Spearman correlation coefficient among replicates (> 0.94).
Randomization	Experimental materials were not divided into random subgroups. Most comparisons were done between WT and mutants.
Blinding	The investigators were not blinded to group allocation during sample collection or analysis, as the information on genotype of materials was essential for the experiment design and analysis.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	FoxA1, Abcam, ab23738, lot#431570, 1:1000 dilution H3, Abcam, ab1791, lot#GR135165, 1:5000 dilution anti-Rabbit IgG HRP, Santa-Cruz, sc2004, #B2216,1:5000 dilution anti-Mouse IgG HRP, Santa-Cruz, sc2005, #B0813. 1:5000 dilution
Validation	FoxA1: Abcam website guarantees the use of ab23738 in WB application for mouse samples. We have demonstrated its specificity in prior work (Iwafuchi-Doi et al., 2016) using FoxA1 knockout cells. H3:Abcam website guarantees the use of ab1791 in WB application for mouse samples.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	H2.35 liver cell line (Zaret, 1998) was derived in house
Authentication	H2.35 cells were authenticated based on the monitoring of morphological features and gene expression patterns by RT-PCR and western blot analyses.
Mycoplasma contamination	Not tested for mycoplasma contamination
Commonly misidentified lines (See ICLAC register)	No commonly misidentified lines were used in this study.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	The heterozygous FoxA2-tRFP knock-in mice were maintained on a C57Bl6/J x CD-1 mixed background. Because none of the E7.5 embryonic phenotypes investigated in this study is evidently sex biased, embryos were not sexed, and a mixed population of male and female embryos (E7.5) was analyzed for RNA-seq and ATAC-seq.
Wild animals	The study did not involve wild animals.
Field-collected samples	The study did not involve field-collected samples
Ethics oversight	All procedures were in accordance with the NIH Guide for the Care and Use of Laboratory Animals and were approved by an IACUC committee at University of Pennsylvania.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Females of 3-5 weeks old FoxA2WT-tagRFP/WT and FoxA2ΔHx-tagRFP/WT were superovulated by intraperitoneal injection of 7.5 IU pregnant mare serum gonadotropin (PMSG) (Prospec, #HOR-272) and, 48 hr later, followed by a second injection of 7.5 IU human chorionic gonadotropin (HCG) (Sigma, #G1063-1VL) and then bred to males of same genotype. E7.5 mouse embryos were dissected into phenol red-free DMEM/F12 (Invitrogen, # 11039-021) supplemented with 5% fetal bovine serum (FBS). Extra-embryonic portion was removed and kept for PCR genotyping. tagRFP intensity of embryos was evaluated by taking a fluorescence image (Nikon TE2000-U) in order to distinguish FoxA2WT-tagRFP and FoxA2ΔHx-tagRFP homozygous embryos from heterozygous and wild type embryos. About 10 embryos of homozygous (for sorting), heterozygous (for setting a FACS gate), or wild type (for setting a FACS gate) were washed with PBS, dissociated with 200 ul of 0.05% Trypsin-EDTA (Invitrogen, #25300054) for 5 minutes at 37 ºC, and then stopped with 200 ul of DMEM/F12 supplemented with 20% FBS. The embryos were pipetted up and down to obtain single cell suspension, spun down to remove supernatant, and re-suspend in DMEM/F12 supplemented with 10% FBS. The cell suspensions were filtered through 35 um filter cap (BD Falcon #352253) and transfer to FACS tube (BD Falcon #352063).
Instrument	BD Influx Cell Sorter
Software	FlowJo, TreeStar, inc
Cell population abundance	The abundance of the cell populations within post-sort fractions are 80.0% for “FoxA2-tagRFP-negative”, 6.45% for “FoxA2-tagRFP-middle”, and 4.10% for FoxA2-tagRFP-high”. We determined the purity by measuring FoxA2 expression level.
Gating strategy	Based on tagRFP intensity of wild type and heterozygous embryos, three sorting gates were set: “FoxA2-tagRFP-negative”, “FoxA2-tagRFP-middle” which were gated by avoiding autofluorescence of WT and including up to maximum tagRFP intensity of heterozygous (FoxA2tagRFP/WT) cells expressed, and “FoxA2-tagRFP-high” which exhibited higher tagRFP signal than heterozygous cells did

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.